Study on Prediction of Top Oil Temperature for Transformers Based on Bayesian Network Model

Ran LI [a], Ya Yuan YI

School of Electrical Engineering, South West Jiaotong University, Chengdu 610031, China

Abstract. The top oil temperature for transformer has a great influence on transformer's operational life and load capacity, therefore,it is important to predict the top oil temperature.On the basis of analyzing and summarizing the main impacts on the top oil temperature,an idea is proposed to predict the top oil temperature by means of Bayesian network,and Bayesian network model is established.The model takes active power,reactive power,load current, ambient temperature and previous time oil temperature as its quantitative indicators, and trains the sample data to find out the probability distribution between various factors.The model is verified according to data collected from the transformer of SSZ11-50kV/220.The results show that the relative error between predictive value and measured value is small,which can be accepted completely in engineering.Therefore,Bayesian network is reasonable and can be widely applied to forecast the top oil temperature.

1 Introduction

With the rapid development of science and technology, electricity that an emerging green energy has been widely used in various fields which range from the national defense science and technology to the people's livelihood. The large demand for electricity lead to the power system developing along the direction[1] of large capacity, high voltage, large grid , power grid,and smart grid,and the development is more and more rapid, which requires power supply and reliability of the power system to achieve a higher level.In power systems,the transformer plays an very important role[2-3] in economic transmission, flexible allocation, safely use of electricity.Winding hot spot temperature is an important performance index of the transformer, and it is closely related to the operation life and the load capacity of transformers.Once transformer's windings get overheated[3],the internal insulation will decreaseand ultimatelyaffect the normal operation of the transformer.However,it is difficult to measure windings hot spot temperature[4].In practice, top oil temperature are usually used insteading of windings hot spot temperature to scale whether the transformer operate normally. Therefore, predicting the top oil temperature timely and taking protective measures in advance notonly to ensure the stable operation of the transformer,but also extendedlife. Which has an important significance in the filed of saving resource and protecting environment.

With constantly update of modern mathematical theory, there are many approaches[5-10]are used to predict transformer's top oil temperature currently,such as

thedirect measurement, numerical analysis, artificial predic- tion and so on.IEEE C57.91-1995 standard [5] has pro-posed transformer top oil temperature rise exponential model, but the model does not fully reflect the impact of ambient temperature.Qing He[6]have taken the load current and ambient temperature into consideration,and establishpredictive models.Comparing withregression model,the predict model is practical.However,but the accuracy remain to be improved.Considering the dynamic changes of ambient temperature around the transformer, B.C.Lesieutre have established an improved top oil temperature prediction model.Model parameters are identified by using the least squares method, and from the predictive results, the model is acceptable in some degree[7].The neural network[8-10] has been applied to predict the top oil temperature according to simulate the relation between the inputs and outputs.

As a branch of mechanical learning,Bayesian network possess the characteristic of inference and visualization when using Bayesian networks to deal with uncertainty and interdependence problem between the multi-factor in complex systems.The generation of oil temperature for transformeris such a problem.There are many factors[11] affecting the top oil temperature,and both of those are interactive.Therefore,Bayesian networks can be used as a effective method of analyzing and forecasting the top oil temperature.After careful analysis and consideration,the paper choose 5 factors,namely active power,reactive power,load current, ambient temperature around transformer and the previous time oil temperature,and build Bayesian

---

[a] LI Ran:1061479141@qq.com

network model according to the 5 parameters.The model has been tested,and obtain predictive data.When predictive datais compared withthemonitor data,it is found that the predictive data is close to monitor data.And when Bayesian network model is compared withBP model,the Bayesian network model outperforms obviously theBP model consistently in relative error.Therefore,Bayesian network model is appropriate to predict transformer's top oil temperature.

2 Descriptionof Bayesian Network

2.1 The Bayesian theory

The Bayesian theory has been proposed by Thomas Bayes in the early 18th century.It is based on Bayesian formula,which is expressed as follows:

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum\limits_{j=1}^{n} P(B_j)P(A|B_j)} \qquad (1)$$

In the formula, B_1, B_2,……,B_n are pairwise mutually exclusive events,and $P(B_i)>0$,i=1,2,3,…,n.Occurrenceof an event "*A*" is always with the simultaneous of one of the B_1, B_2,…,B_n.Bayesian formula is to find the probability of each event "B_i" lead to "*A*" under the condition that the event "*A*" has happened.The top oil temperature is affected by many factors.In the processing of predicting the top oil temperature,the probabilities of the events that affect the top oil temperature is calculated by Bayesian formula,and then to forecast transformer's oil temperature by means of the probabilities.

2.2 The Bayesian network

Bayesian network has been proposed by Judea Pearl in 1988. Its essence is an inference network,which is based on uncertainty probability. Bayesian networkprovides an effective method with regard to the causal information presentation.In addition, it is also one of the most effective theoretical models in the areas of expression and reasoning of uncertain knowledge currently[12].Bayesian network is a directed acyclic graph (DAG), that is made up of nodes and directed edges.Nodes represent random variables,and directed edges between nodes represent the relationship between the nodes (points to his son by the parent node). The intensity of the relationship between nodes is expressed by the conditional probability,If there is no parent nodes,it is expressed by the prior probability.
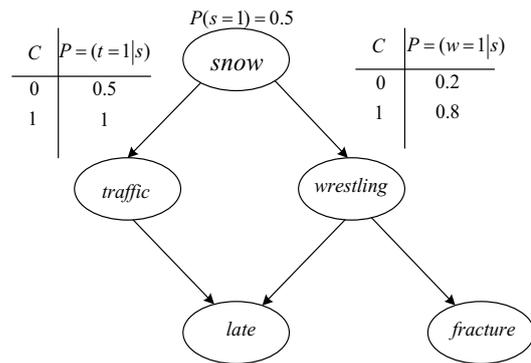
The following Figure1 is a simple Bayesian network.

Figure1.a simple Bayesian network

In the Figure1,there are 5 random variables,namely snow, traffic jam, late,wrestling and fracture,and they are briefly represented by "s", "t", "l", "w"and "f".Variables of 1 or 0, is on behalf of the event is true or false.For example, the probability of snowing is 0.5,which can be expressed by P(*s*=1)=0.5.And P(*t*=1|*s*) represents that the probability of traffic jam is 1 if snowing.According to causality,the *s* is a parent node of *w*,and f is a child node of *w*.

2.3 establishment of Bayesian network model

In order to predict the top oil temperature by the Bayesian network,the first step is to establish Bayesian network model.The process of establishing model is divided into 3 steps:
1) choosing the variables nodes
2) determining the network structure
3) deciding the probability distribution of variables

In order to predict the top oil temperature precisely,it is important to take the key factors that influence the top oil temperature into account.However,it is impossible to think over allfactors.The paper divides elements into 3 sorts:environmental parameters,characteristicparameters, and running parameters.Meanwhile,those parameters are interactive.For example,the relationships exiting in the power,voltage,and currentcan be describe as follows:

$$P = UI\cos\varphi \qquad (2)$$

$$Q = UI\sin\varphi \qquad (3)$$

Where *P* is active power,*Q* is reactive power,*U* is voltage,and *I* is load current.*φ* is phase angle.

After consideration,active power,reactive power,load current, ambient temperature around the transformer and historical oil temperature are taken as the variables nodes.

Generally speaking,there are 3 ways to build network structure[13].The first method is that the experts in this field determine network nodes and structure, and provide distribution parameters of the node variables .The second one is that network nodes and structure are determined by experts,butacquire distribution parameters fromsample data by machine learning.The third one is that experts only identify network's nodes, and its structure and parameters are acquired by machine learning.The first method relies entirely on expert knowledge, and expert cognitive is limited, so prediction results will appear large deviation due to the limitations.The first method is just suitable for some simple questions. The second method is fit for the

prediction problem where relationship between variables is clear and there are large capacity of sample data.The efficiency of forecast analysis can be improved through parameters learning.The third way is completely dependent on data-driven,so it is apt to address the questions that have larger variable quantity, more complex relationship and adequate learning samples.The prediction of top oil temperature belongs to the problems are described by the above third method.So the network structure can be decide easily.

After determining the network structure,the next step is to calculate the conditional probability of variables.In Bayesian networks,the probability distribution can be obtained by training the sample data.But the sample data must have a large number.The one of the most principal steps is network inference[17],by which the predictive results can be obtained.

3 Bayesian probability distribution calculation

3.1 compute of condition probability

Theconditional probabilitieswhichreflectthe dependent relationships of variousfactors，it plays very important roles in the emergence of top oil temperature.Only find out the conditional probabilities,can predict the top oil temperature.At present,there are a lot of algorithms[14-15]applied to compute conditional probabilities.such as the least squares(LS),moment method(MM),and maximumlikelihood estimate (MLE).This paper selects MLE.

Assume sample is $O=\{O_1,O_2,......O_m\}$,the likelihood function [16]can be expressed by:

$$L=\sum_{i=1}^{n}\sum_{m=1}^{M}\log P(X_i|P_a(X_i),O_m) \tag{4}$$

When the network node variable is discrete,there will be $\theta_{i,j,k}=P(X_i=k|P_a(X_i)=j)$,so the formula(4) can rewrite as

$$L=\sum_i\sum_m\log\prod_{j,k}\theta_{i,j,k}^{I_{i,j,k,m}}=\sum_i\sum_m\sum_{j,k}I_{i,j,k,m}\log\theta_{i,j,k}=$$
$$\sum_{i,j,k}I(X_i=k,P_a(X_i)=j|O_m)\log\theta_{i,j,k} \tag{5}$$

$$I_{i,j,k,m}=\begin{cases}0, & X_i=k,P_a(X_i)=j\\1, & others\end{cases} \tag{6}$$

Where, $\sum_m I(X_i=k,P_a(X_i)=j|O_m)$ recordsthe numberof $(X_i=k,P_a(X_i)=j)$ .By Lagrange multiplier method,the maximum likelihood estimate of $\theta_{i,j,k}$ can be calculated, the result is

$$\hat{\theta}_{i,j,k}=\frac{\sum_m I(X_i=k,P_a(X_i)=j|O_m)}{\sum_m\sum_k I(X_i=k,P_a(X_i)=j|O_m)} \tag{7}$$

3.2 Bayesian network inference

Bayesian networkinference is to calculate probability of nodes variable by means of Bayesian network structure and conditional probability table after giving sample data . Probability inference is one of the most key steps in the process of forecasting top oil temperature.In the past 20 years,the Bayesian network has developed some effective inference algorithms,that consisting of accurate algorithm and approximatealgorithm.The accurate algorithm makes conclusions in full accordance with the calculation formula,so it is just suitable to addressing simple questions.But approximate algorithm is widely used in large-scale network structure to solve the problem,that is complex and have a large amount of calculation.The prediction of the top oil temperature involves a lot of variables,and the variables are interactive,so it is an extremely complicated question. Therefore it is necessary to adopt approximate algorithm.The commonly used approximate algorithm in network inferenceinclude stochasticsamplingalgorithm, search-basedalgorithm,model simplification algorithm, and loopy belief propagation algorithm.The paper makes use of stochastic sampling algorithm.

4 Prediction of top oil temperature by Bayesian network

The paper uses Bayesian Network Tool(BNT) in Matlab to realize the establishment of network structure, network variable parameter learning and prediction of top oil temperature.Matlab is a strong tool in numeral calculations, It can not only reduce computational expense greatly,but also improve the prediction efficiency in some degree.

4.1 sample data

The data used in the paper is provided by a certain grid company in Chengdu.Sampling a transformer of SSZ11-50kV/220 every five minutes from 0:00 to 24:00 on May 3 and 4,we can easily obtain top oil temperature data.And ambient temperature around the transformer,load current, active power and reactive power of the corresponding moment are also supplied. The sample data is divided into training data and test data according to the ratio of 9:1.So there are 520 sets of data used to training Bayesian network,56 sets of data used to test the correctness of model.

4.2 training and learning of Bayesian network

Bayesian network structure can be acquired from learning sample data.There are mainly two types of commonly use structure learning method, namely based on dependency test study and based on search score learning.The latter is widely used,and its principle is to establish Bayesian network structure according to certain search strategy and score criterion in all nodes structure space.This paper has adopted the based on search score learning.Constructing Bayesian network structure consists of two steps:model selection and model optimization.K2 algorithm is applied to learn Bayesian network structure in the paper.It uses posteriori probability as scoring function. According to the pre-determined network node sequence, the posteriori

probability associated with the node is obtained by using the greedy search,and decide the node of largest posterior probability as its parent.After allparent nodes determined, the construction of Bayes networkstructure is completed finally.Then the following step is to calculate conditional probability.Above 2.1,maximum likelihood estimate has been introduced in detail.

4.3 prediction results and analysis

At present,there are many indicators are used in most of literature to represent the deviation of the predictive values and actual values,such as squared error and mean absolute error, mean relative error and meansquare error. The paper select relative error to evaluate the prediction accuracy.

56 sets of data are used to test Bayesian network,and 56 predictive top oil temperature are generated,and the comparison with actual data is expressed in the Figure2 BP neural network also apply to prediction of transformer top oil temperature,and the predicted values are compared with the actual values.The comparison is described in the following Figure3:
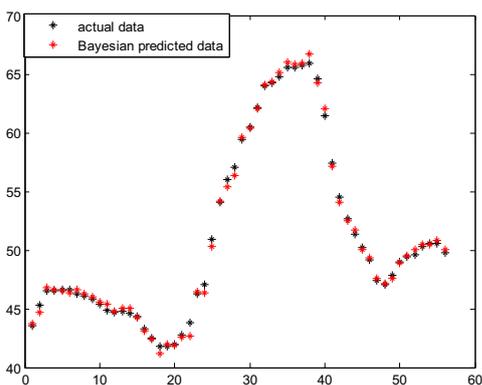

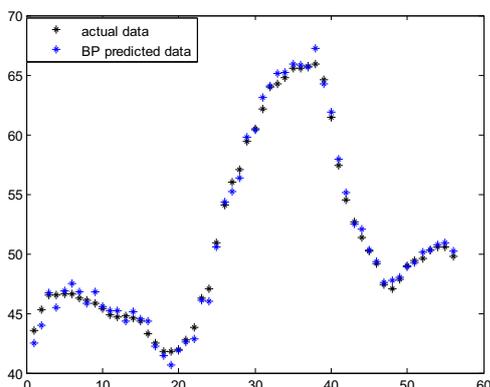**Figure 2.**The waveform of Bayesian predicted value and actual value

Figure 3. The waveform of BP predicted value and actual value

From Figure2 and 3,we can easily have a simple cognition that Bayesian forecasting result is closer to actual value.

In order to observe results clearly,the relative error

between predictive values and measured values have been computed,and I paint the results in a picture,which is represented in Figure4.
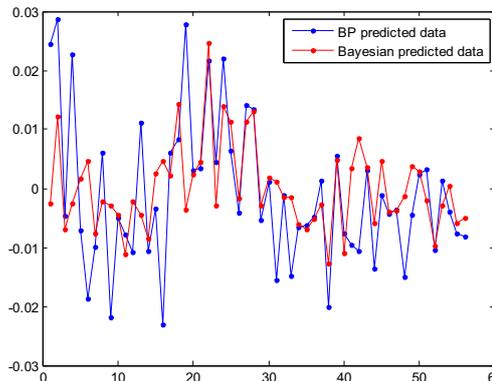

Figure 4. The waveform of relative between predicted value and actual value

As we can see,the relative error of Bayesian network fluctuate around"zero",and the fluctuationranges small. When the Bayesian network is compared with the BP model, the Bayesian network is superiorto theBP model consistently in relative error part.The Bayesian network's maximum relative error is 1.82%,which can be accepted absolutely in engineering.The predicted data can be used to assist decision-maker to guide transformer operation normal.

5 Conclusions

This paper analyzes some factors affecting transformer top oil temperature, and determines the five main factors as Bayesian networks' random variables.On the basis of the analysis of Bayesian network, sampling information active power, reactive power, ambient temperature, load current and the last time the top oil temperature are used as the inputs to train Bayesian network by means of K2 algorithm.And the optimal network structure is determine as a result of employing greedy search.And finally,taking advantage of stochastic sampling algorithm to deduce the transformer top oil temperature.In the end,the prediction of top oil temperature for transformer is realized.Due to many factors that affect the oil temperature, and this article only selects five variables, so deviation exist in forecast results.But the error is acceptable to some extent. Therefore,proposing Bayesian network to predict top oil temperature for transformer can be applied to practice.

References

1. WANG Yongqiang, YUE Guoliang, HE Jie, et al. Studyon Prediction of TopOil Temperaturefor PowerTransformerBasedon Kalman Filter Algorithm [J]. High Voltage Apparatus,**50(8)**,74-79 (2014)
2. XIONG Hao, CHEN Weigen, DU Lin, et al. Study on prediction of Top-oil Temperature for Power Transformer Based on T-S Model[J]. Proceedings of the CSEE,**27(30)**,15-19 (2007)

3. CHEN Jinming, WU Yi，ZHU Haibing, et al. Study and application of Transformer Top-oil predicting[J]. Electrotechnical Application, **33(22)**,89-93 (2015)

4. LIU Hui,SONG Guobing,LUO Junming.Dynamic modeling based on the transformer top oil temperature[J].TRANSFORMER,**39（9)**,5-9(2002)

5. IEEE Standard C.57.91 — 1995 IEEE guide for loading mineral-oil-immersed power transformer[S](1995)

6. Qing H，Si J, Tylavsky J D. Prediction of top-oil temperaturefor transformers using neural networks [J]．IEEE Transactions on Power Delivery, **15(4)**，1205-1211 (2000)

7. Lesieutre B C，Hagman W H，Kirtley Jr J L．An improved transformer top oil temperature model for use inan on-line monitoring an diagnostic system [J]．IEEE Transactions on Power Delivery,**2(1)**，249-256 (1997)

8. Galdi V,Ippolito L，PiccoloA，et al.Neural diagnostic system for transformer thermal overload protection[J]．Proc. Inst. Elect. Eng.，Elect. Power Application，**147(5)**,415-421 (2000)

9. Pradhan M K，Ramu T S．On-line monitoring of temperature in power transformers using optimal linear combination of ANNs[C]．IEEE International Symposium on Electrical Insulation，Indianapolis，USA，70-74 (2004)

10. CHEN Weigen, XI Hongjuan, Su Xiaoping, et al. Application of Generalized Regression Neural Network to Transformer Winding Hot Spot Temperature Forecasting[J]. High Voltage Engineering, **38(1)**,16-21 (2012)

11. CHEN Weigen, SU Xiaoping, CHEN Xi, et al. Influence Factor Analysis and Improvement of the Thermal Model for Predicting Transformer Top Oil Temperature[J].High Voltage Engineering , **37(6)** , 1329-1335(2011)

12. HU Yusheng,TU Xuyan,CUI Xiaoyu,et al. Ainferential Method of Uncertain Knowledge Based onBayes-Network[J].ComputerIntegrated manufacturing System,**7(12)** , 65-68 (2001)

13. ZHOUZhiyuan, SHEN Guchao,ZHU Xiaolong.et al. Application of Bayesian Network in Intelligence Prediction[J].Information Science,**32(10)**,3-8(2014)

14. Wang YansongYu Jilai．Joint conditions probability forecastmethod for wind speed and wind power[J]．Proceedings of the CSEE ，**31(7)**,7-14 (2011)

15. REN Jia，GAO Xiaoguang.*Bayesian network parameter learning anddecisionsupport forUAVs*[M]. Beijing：National Defence Industry Publisher，70-100 (2012)

16. JANGFeng,GAOWen,YAO Hongxu.The *inference andlearning ofBayesian networks*[C]// Proceedings of national networkand informationsecuritytechnology conference in 2005.268-275(2005)

17. LI Haitao,JIN Guang,ZHOU Jinglun,et al.Survey of Bayesian network inference algorithms[J].Systems Engineering and Electronics,**30(5)**,935-939 (2008)